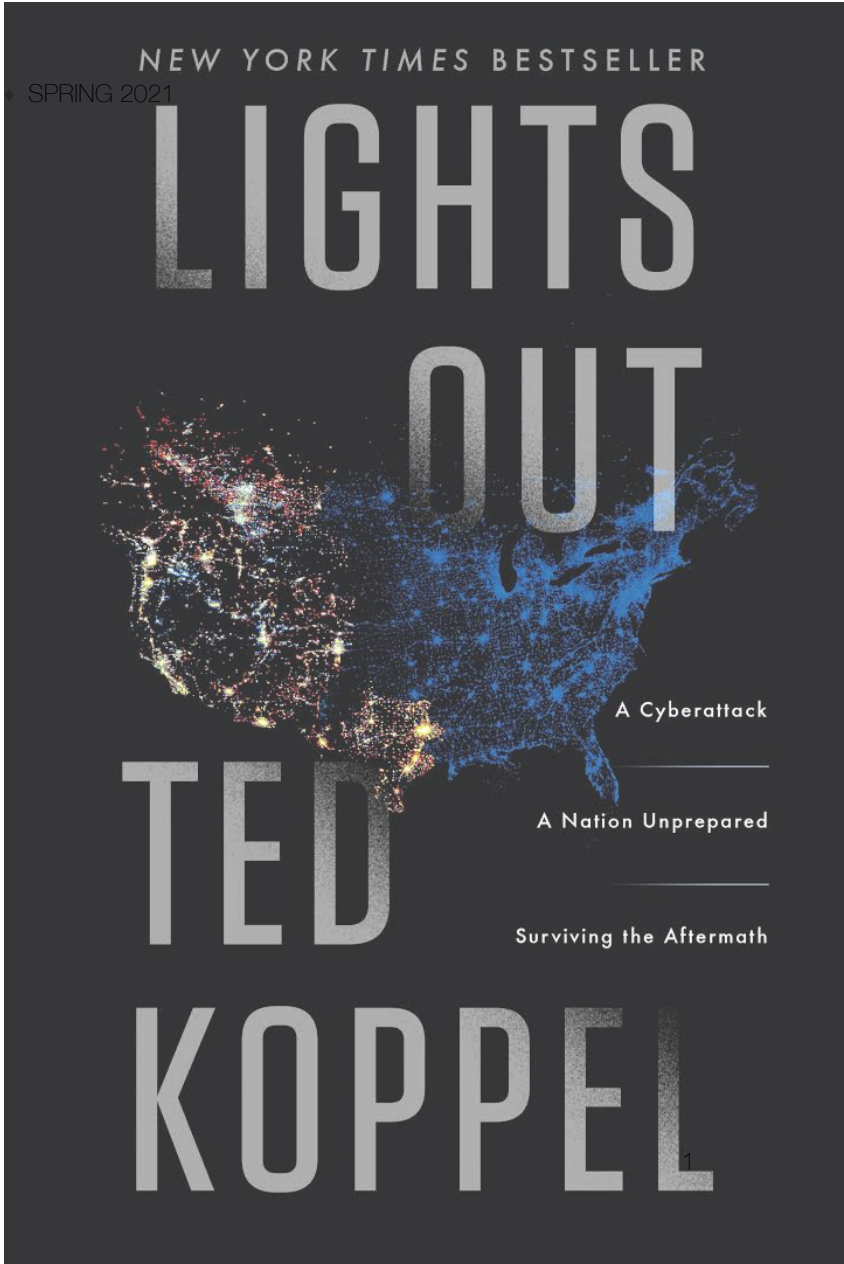
TERM PROJECT

Critical Infrastructure Protection

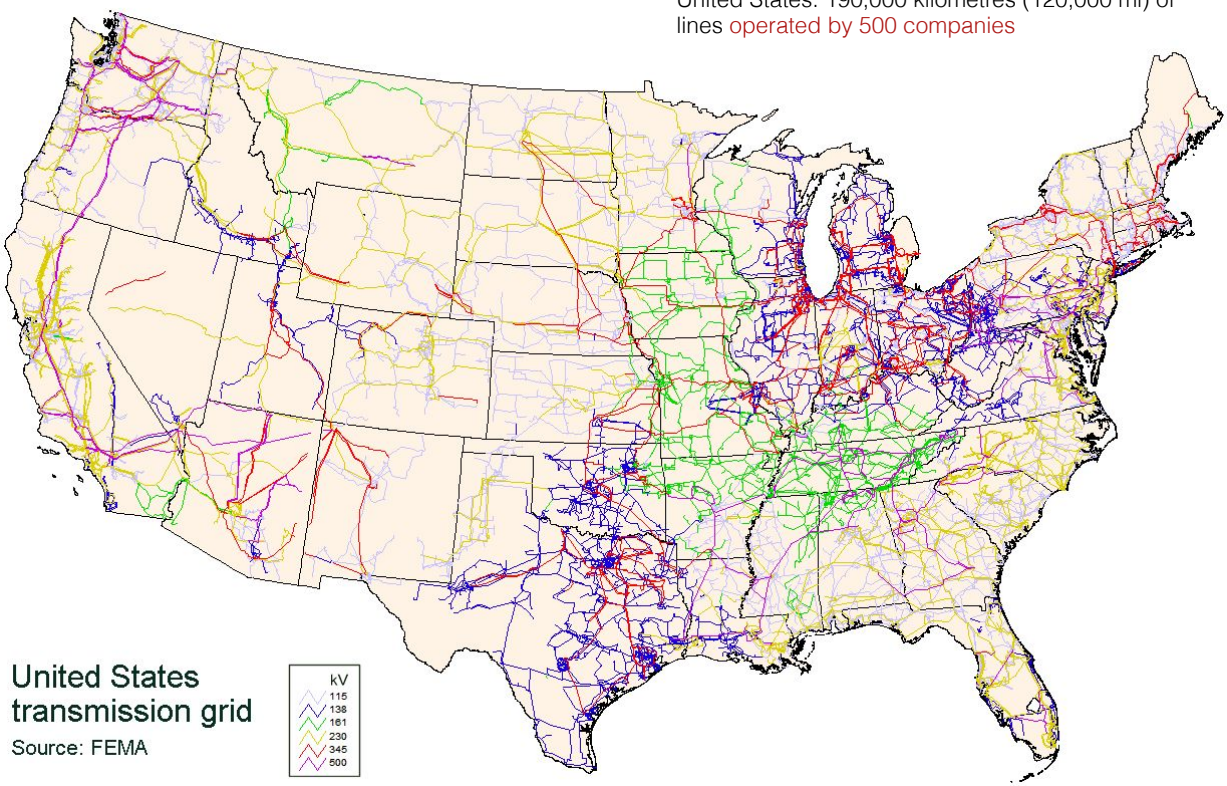
"So many of our transactions are now conducted in cyberspace that we have developed dependencies we could not even have imagined a generation ago. To be dependent is to be vulnerable. We have grown cheerfully dependent on the benefits of our online transactions, even as we observe the growth of cyber crime. We remain largely oblivious to the potential catastrophe of a well-targeted cyberattack."

— Ted Koppel, 2015



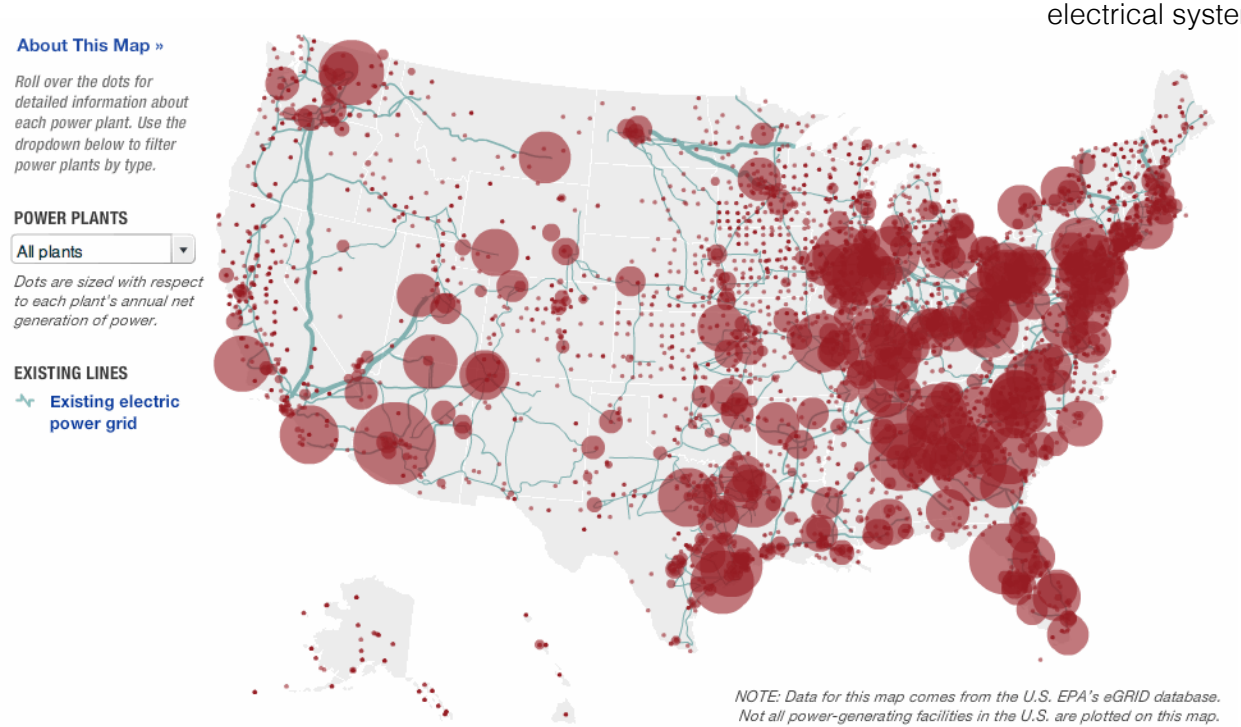
Electric Power Grid

Electric power transmission grid of the contiguous United States: 190,000 kilometres (120,000 mi) of lines operated by 500 companies

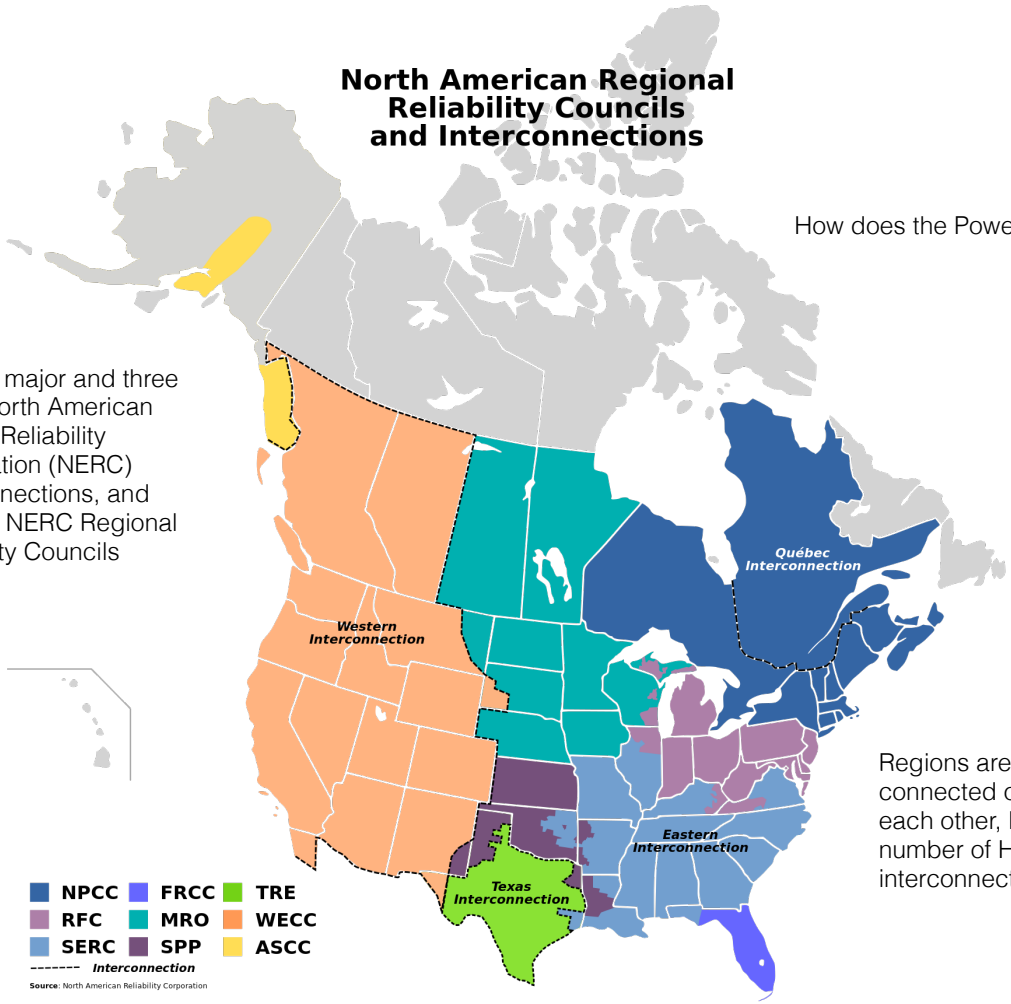


Security Problems

"The U.S. electric grid is a complex network of independently owned and operated power plants and transmission lines. Aging infrastructure, combined with a rise in domestic electricity consumption, has forced experts to critically examine the status and health of the nation's electrical systems."



Source: https://www.infrastructureusa.org/interactive-map-visualizing-the-us-electric-grid/



How does the Power Grid work? (video)

The two major and three minor North American Electric Reliability Corporation (NERC) interconnections, and the nine NERC Regional Reliability Councils

Regions are not directly connected or synchronized to each other, but there are a number of HVDC interconnections.

The U.S. power grid consists of three loosely connected parts, referred to as interconnections: Eastern, Western and Texas.

Within each, **high-voltage power lines transmit electricity** from generating sources such as coal or hydroelectric plants to local utilities that distribute power to homes and businesses, where lights, refrigerators, computers and myriad other “loads” tap that energy.

Because **electricity in power lines cannot be stored**, generation and load have to match up at all times or the grid enters blackout territory. The **interconnectedness** of the grid makes it easier to compensate for local variations in load and generation but it also gives blackouts a wider channel over which to spread.

Transmission system operators scattered across some 300 control centers nationwide monitor voltage and current data from **SCADA (supervisory control and data acquisition)** systems placed at transformers, generators and other critical points.

Source: Scientific American
https://www.scientificamerican.com/article/2003-blackout-five-years-later/

The 2003 Northeast Blackout

On August 14, 2003, shortly after 2 P.M. Eastern Daylight Time, a high-voltage power line in northern Ohio brushed against some overgrown trees and shut down—a fault, as it's known in the power industry. The line had softened under the heat of the high current coursing through it.

Normally, the problem would have tripped an alarm in the control room of FirstEnergy Corporation, an Ohio-based utility company, but the alarm system failed.

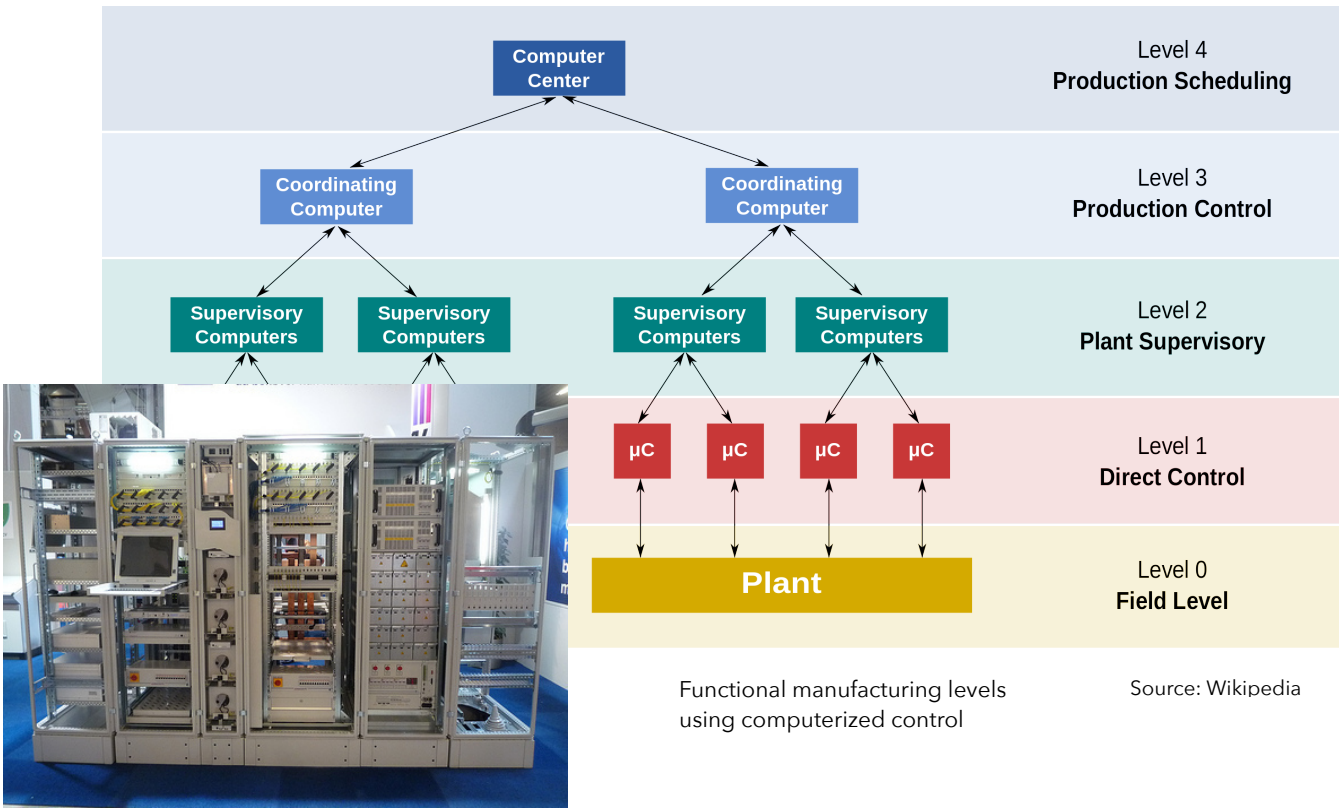
Over the next hour and a half, as system operators tried to understand what was happening, three other lines sagged into trees and switched off, forcing other power lines to shoulder an extra burden. Overtaxed, they cut out by 4:05 P.M., tripping a cascade of failures throughout southeastern Canada and eight northeastern states.

50 million people lost power for up to two days in the biggest blackout in North American history. The event contributed to at least 11 deaths and cost an estimated $6 billion.

Source: Scientific American
https://www.scientificamerican.com/article/2003-blackout-five-years-later/

SCADA Systems

Supervisory control and data acquisition (SCADA) is a **control system architecture**



Functional manufacturing levels using computerized control

Source: Wikipedia

Electric Grid Security (video)

"These days, within any one of the three U.S. grids, almost all operational phases of thousands of power companies are interconnected. Coordinating operations are run using the same supervisory control and data acquisition (SCADA) systems. Most of the systems are manufactured by a relative handful of companies, and while they are not quite interchangeable, there are similarities in programming and structure. This presents a web of pathways connecting the thousands of power companies and enabling transactions … The overall system has been designed for maximum efficiency, eliminating waste while establishing a precise balance between the power needed and the power generated."

(video)

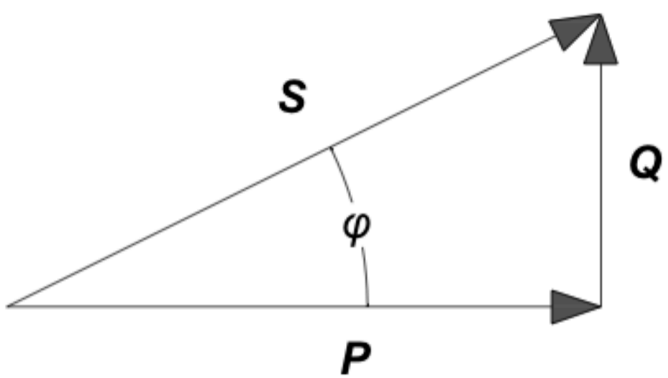
Active, Reactive and Apparent Power

Real, imaginary and apparent power consumption in AC circuits

The required power supply to an electric circuit depends on the

- **active power** (P) - real electrical resistance power consumption in circuits
- **reactive power** (Q) - imaginary inductive and capacitive power consumption in circuits

The required power supply is called the **apparent power** (S) and is a complex value that can be expressed in a Pythagorean triangle relationship as indicated in the figure below.



Apparent Power - S

The apparent power is the power supplied to the electric circuit—typical from a power supplier to the grid—to cover the real and reactive power consumption in the loads.

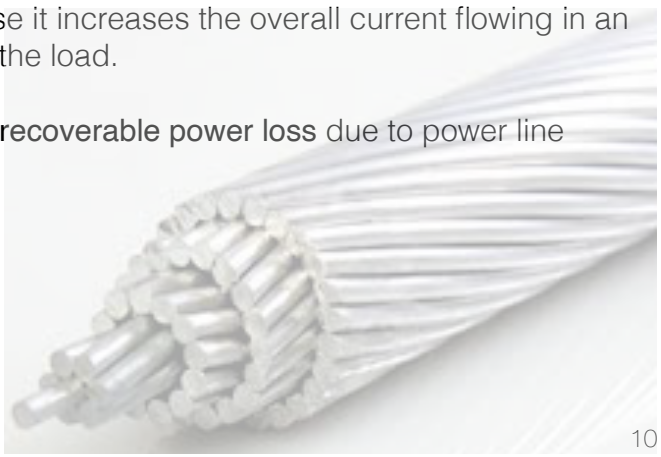
Active Power - P

Active—or real or true—power does the actual work in the load. Active power is measured in watts (W) and is the power consumed by electrical resistance.

Reactive Power - Q

Reactive power is the imaginary or complex power in a capacitive or inductive load. Reactive power represents an **energy exchange between the power source and the reactive load**s where no net power is gained or lost. The net average reactive power is zero. Reactive power is stored in and discharged by inductive motors, transformers, solenoids and capacitors.

- Reactive power should be minimized because it increases the overall current flowing in an electric circuit without providing any work to the load.
- Increased reactive currents only provides **unrecoverable power loss** due to power line resistance.



Term Project Description

Our term project is a group project comprising **three separate parts** as detailed below. Each group has three members working together as a project team.

Project Deliverables

1. Developing predictive models for behavioural anomaly based online intrusion detection based on monitoring and analyzing control signals streamed in real time from the continuous operation of a cyber-physical system

2. **Project report** is to be completed and submitted by **APRIL 9, 2021**. Detailed description of tasks will follow.

3. **Presentation slides** to be submitted by **APRIL 12, 2021**.

4. Presentations of project outcomes on **APRIL 13, 14** and **16**: note that in addition to the regular class hours we need extra time slots to accommodate all presentations. Please make your reservations early (first come, first served).

Project Scope

- In light of ever increasing cyber threats, especially advanced persistent threats, and existing vulnerabilities that expose critical infrastructure to a variety of adversarial scenarios, the project explores behavioral anomaly based intrusion detection methods used for cyber situational awareness of automated control processes.
- …

Project Scope

- …
- Electric power grids, intelligent transportation systems, public water utilities, oil and gas pipelines, and other critical infrastructure routinely rely on automated control for their **continuous operation**. Automation is essential for operating equipment and monitoring conditions of machinery, production processes and plants; it enhances efficiency and quality of service delivery, safe operation of critical assets and their protection in case of internal or external disruptions.
- …

Project Scope

- …
- …
- Supervisory control is essential for cyber situational awareness. Real-time control data from the continuous operation of a **cyber-physical system**[1] is **stream data**. Technically, such data forms a *multivariate time series* that needs to be analyzed in (near) real-time. Continuous data analysis is vital for early threat detection—suspicious, i.e. potentially harmful, behavioral anomalies—to mitigate the impact of attacks by launching countermeasures to contain the damage and facilitate cyber forensics.

> [1]*Cyber-physical system*: "Embedded computers and networks monitor and control the physical processes, with feedback loops where physical processes affect computations and vice versa. The technology builds on the discipline of embedded systems, computers and software embedded in devices whose principle mission is not computation, such as cars, toys, medical devices, and scientific instruments. CPS integrates the dynamics of the physical processes with those of the software and networking, providing abstractions and modeling, design, and analysis techniques for the integrated whole." (cyberphysicalsystems.org)

Challenges

A number of inescapable ‘external factors’ can make anomaly detection in time series data challenging whenever data originates from the operation of a real-world system. Typical examples include:

- **imperfections in the data**, such as missing or corrupted values;
- lack of ground truth in historic data;
- unavailability of labels to differentiate normal data points from outliers;
- **various types of anomalies** depending on the application context;
- striking a good balance between precision and recall, specifically also
- reducing the false alarm rate to make anomaly detection practical in any real application context with resource constraints.

Getting Started

You will receive additional input verbally and in writing as you proceed through your project. A clear breakdown of tasks and responsibilities shared by the team members helps in developing a CLEAR ROADMAP allowing the team to work more productively.

Please take into account that **the team as a whole** is responsible for their project; that is, team members are expected to help each other in managing project tasks—just like in a real professional environment.

We hope you will find this project an interesting and rewarding experience.

Thank you for your cooperation!